

UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF MISSOURI
EASTERN DIVISION

ADDINGTON STEWART, et al.,)	
)	
Plaintiffs,)	
)	
vs.)	Case No. 4:04CV00885 RWS
)	
CITY OF ST. LOUIS,)	
)	
Defendant.)	
)	
and)	
)	
PAUL E. DAVIS, et al.,)	
)	
Intervenors.)	

MEMORANDUM AND ORDER

This matter is before me for a ruling after a bench trial.

Plaintiffs Firefighters' Institute for Racial Equality and four individual African-American firefighters (collectively "FIRE" or the "FIRE plaintiffs") have filed this action against the City of St. Louis (the "City") under Title VII of the Civil Rights Act of 1964, 42 U.S.C. § 2000e, et seq., and 42 U.S.C. §§ 1981 & 1981a. The FIRE plaintiffs claim that the promotion tests for the City of St. Louis Fire Department in 2000 and 2004 for Fire Captain and Battalion Chief had an adverse impact on African-Americans, and seek injunctive relief and damages.

The St. Louis Fire Fighters' Association International Association of Fire Fighters Local 73 and several Caucasian firefighters (collectively "Local 73") are intervenors in this case. The intervening Caucasian firefighters took the 2000 and 2004 promotion tests and are on the promotion lists challenged in this case.

Prior to trial I granted summary judgment in favor of the City and Local 73 on the issue of whether the City *intentionally* discriminated against Plaintiffs in the testing process.

Based on the evidence presented at trial, I find that the 2000 Fire Captain examination did have an adverse impact on African-Americans. However, I find that the 2004 Fire Captain examination and the 2000 and 2004 Battalion Chief examinations did not have an adverse impact on African-Americans. Alternatively, even if all of the tests could be determined to have had an adverse impact on African-Americans, I find that the tests were “job-related for the position in question and consistent with business necessity,” as required by 42 U.S.C. § 2000e(k), because they tested the relevant knowledge, skills and abilities that underlie the performance of the Fire Captain and Battalion Chief positions. The FIRE plaintiffs did not present any evidence that an alternative selection method was available that has a less disparate impact and would still serve the City’s legitimate business interest.

I will enter judgment in favor of the Defendant City of St. Louis and Intervenors.

I. BACKGROUND

A. History of Recent Litigation

The relevant factual background of this case began in 1991 when, as recounted in St. Louis Firefighters Ass’n v. City, 96 F.3d 323 (8th Cir. 1996), the City announced plans to implement a testing procedure for the promotion of Fire Captains to Battalion Chiefs in the City Fire Department. The City devised a three-part test, consisting of a written test, a fire scene test, and an assessment center test. The test battery was structured so that candidates had to pass each test in order to proceed to the next. The written test was administered in February 1992 to 54 candidates, 20 of whom were African-American. The fire scene test was administered in June

1993 to 34 candidates, 10 of whom were African-American. The assessment center test was administered in April 1994 to 23 candidates, 7 of whom were African-American.

In June 1993, after considering several schemes for weighting the multiphased test results, and after both Local 73 and FIRE objected to various formulations, the City proposed to make the written and fire scene tests pass/fail, and then to rank passing candidates solely according to their assessment center scores. Of the 23 candidates who took the assessment center test, African-Americans ranked eighth, eleventh, twelfth, seventeenth, eighteenth, twenty-second, and twenty-third. Upon considering these results, and contemplating possible litigation on the basis of an adverse impact against African-Americans, the City reweighted the test battery results. While keeping the written test pass/fail, the City made the fire scene test worth 30% of the ranking score, and made the assessment center worth 70% of the ranking score. Following this reweighting, white candidates tended to drop in the ranking, and African-American candidates, who now ranked third, fifth, eighth, eleventh, twelfth, eighteenth, and twenty-third, tended to improve in ranking. Local 73 and FIRE both filed complaints against the City seeking to bar the City from making promotions based on the test procedures. Before trial, the City canceled the entire promotion test procedure and announced its intention to establish a new testing procedure, and the District Court dismissed the action as moot. *Id.*¹

In 1998, FIRE and twenty-two individual plaintiffs filed an action against the City of St. Louis under 42 U.S.C. §§ 1981 and 1988, Title VII of the Civil Rights Act of 1964, 42 U.S.C. § 2000e, and the Missouri Human Rights Act, Mo. Rev. Stat. § 230.010, *et seq.*² The plaintiffs

¹*See Firefighters Institute for Racial Equality v. City of St. Louis*, 616 F.2d 350 (8th Cir. 1980), for a history of earlier promotional practices of the St. Louis Fire Department.

²*Firefighter's Institute for Racial Equality, et al. v. City of St. Louis*, 4:98CV00789 Slip Op. (E.D. Mo., Sept. 2, 1999).

claimed that the 1997 test for the promotion of Fire Captains to Battalion Chiefs in the City's fire department was "illegal" because the test had a disparate impact on African-Americans. The individual plaintiffs were African-American Fire Captains who applied for but were denied promotion to Battalion Chief. On summary judgment, the District Court concluded that the evidence, viewed in the light most favorable to the plaintiffs, failed to present a genuine issue of material fact regarding the validity and job-relatedness of the employment tests and process used by the City for the promotion of Fire Captains to Battalion Chiefs in the City's fire department. Accordingly, the District Court held that the City was entitled to judgment as a matter of law on the plaintiffs' Title VII disparate impact claim, a holding affirmed by the U.S. Court of Appeals for the Eighth Circuit in Firefighters' Inst. for Racial Equal. v. City of St. Louis, 220 F.3d 898 (8th Cir. 2000).

B. This Case

The evidence presented in this case established the following: The City conducted promotion tests for Fire Captain and Battalion Chief in 2000, 2002,³ and 2004. Plaintiffs Leonard Davis and Dominguez Jones took the 2000 and 2004 Fire Captain promotion tests, but neither Davis nor Jones was placed on the promotion list for Fire Captain. Plaintiffs Addington Stewart and Michael Richardson took the 2000 and 2004 Battalion Chief promotion tests, but neither Stewart nor Richardson was placed on the promotion list for Battalion Chief. The 2000

³Local 73 and several Caucasian firefighters filed a lawsuit in this Court challenging the 2002 promotion tests before the tests were given by the City. In that case, I ruled against Local 73 and dismissed the Complaint because it was not ripe for adjudication. Barnes v. City of St. Louis, Mo., 4:02CV25, Slip Op. at 1 (E.D. Mo. Feb. 19, 2002). The City later abandoned the 2002 test results for reasons unrelated to the Complaint.

promotion tests for Fire Captain and Battalion Chief were designed and administered by Barrett & Associates. The 2004 promotion tests were designed and administered by a firm named SHL.

The FIRE plaintiffs filed an EEOC Charge of Discrimination alleging that the 2000 tests did not test the relevant skills and abilities of the examinees and that the tests had an adverse impact on African-Americans. On August 22, 2003, the EEOC issued its finding that both the 2000 Battalion Chief and Fire Captain tests had an adverse impact on African-Americans, that the City failed to establish content validity for the 2000 Fire Captain test under the Uniform Guidelines for Employee Selection Procedures (“UGESP”), and that the 2000 Battalion Chief test “did not test the skills and abilities for that position and thus did not meet the [UGESP].”

The EEOC attempted to conciliate the charge without success.

Plaintiffs filed an EEOC Charge of Discrimination on March 25, 2004, alleging that the 2004 tests “were discriminatory” to African-Americans because:

- 1) A white female firefighter captain was given an extra hour to complete the written portion of the promotional exam;
- 2) The grading system used by the testing firm (SHL) does not meet the EEOC guidelines regarding testing;
- 3) The weights given to the parts of the examination do not reflect the jobs being tested for, as required by the guidelines;
- 4) The service rating component of the qualifications for promotion is biased and prevents African Americans from getting a fair opportunity to be promoted;
- 5) The grading system used by SHL does not meet the federal guidelines for testing. For example, with regard to the oral portions of the exams, the candidates performances were not taped, so that it will be impossible to verify, correlate or validate the scores of the individual raters – a violation of the EEOC guidelines;

- 6) Several white firefighters who took the examinations dressed in dark suits to alert the assessors that they were union representatives;
- 7) The use of this new promotional testing adversely impacts those African American firefighters who took the promotional exams in 2000 and were found to have been discriminated against in the 2000 testing process;
- 8) The company hired to administer the 2004 promotional exams, SHL, is merely a reformulation of the firm Landy Jacobs, which previously engaged in exam impropriety in St. Louis.

In early July 2004, while the 2004 EEOC charge was pending, the City announced its intention to promote individuals from the 2004 promotion lists. The FIRE plaintiffs filed this lawsuit on July 16, 2004.

II. FINDINGS OF FACT

A. 2000 Fire Captain and Battalion Chief Tests

In 1999, the City retained the firm of Barrett & Associates to develop and administer promotional examinations for the positions of Fire Captain and Fire Battalion Chief in the St. Louis Fire Department. Gerald V. Barrett, Ph.D., the principal of Barrett & Associates (“B&A”), together with Michael Polomsky, Ph.D., a B&A employee who served as the project manager, and Winfred Arthur, Ph.D., a B&A subcontractor, were the individuals principally involved in the development and administration of the 2000 Fire Captain and Battalion Chief examinations. B&A began its work by performing job analyses for the positions of Fire Captain and Battalion Chief using archival material, such as current and previous job descriptions and job interviews with incumbents who were promoted from the 1997 tests for Battalion Chief, as well as semi-structured interviews with present job incumbents.

Dr. Winfred Arthur was in charge of the job analyses and, ultimately, the production of comprehensive job descriptions for each position. Dr. Arthur was assisted in his interviews of job incumbents by Dr. Barrett and Dr. Polomsky. Based on his job analyses, Dr. Arthur created job descriptions which listed the major work behaviors for each position in descending order of importance and also listed the knowledge, skills and abilities (“KSAs”) necessary to perform each major work behavior. Before the job descriptions were finalized, the job descriptions were returned to the City for the purpose of review by the incumbents who had been previously interviewed for their comments and proposed modifications.

On the basis of the job descriptions, source materials, also known as reading lists, were identified. Reading lists were made available to the candidates well in advance of the administration of the tests. Test “items” or questions were also developed on the basis of the job descriptions and the reading lists derived therefrom.

There were three components to the 2000 Fire Captain examination: 1) a 100-question job knowledge multiple choice test, with three alternative answer choices for each question, 2) a 50-item “write- in / mark-in” fire scene simulation with 14 multiple choice questions, and 3) an oral briefing exercise. There were four components to the 2000 Battalion Chief examination: 1) a 100-question job knowledge multiple choice test, with three alternative answer choices for each question, 2) a 50-item “write- in / mark-in” fire scene simulation, 3) an oral presentation exercise, and 4) a 50-question in-basket exercise designed to test the administrative components of the Battalion Chief’s job description, which candidates were given three hours to complete.

Whenever an item is written for a test, the writer is required to justify why what they're proposing as the best answer is, in fact, the best answer. As a result, the writer has to document a justification for each right answer. Dr. Winifred Arthur testified that “this typically will entail

going back to the source material and identifying the section of the source material that says this is the way you're supposed to do this.” These “justifications” were created for all components of both tests prior to the administration of the tests. Additionally, each test item underwent an “iterative review process” by which, before inclusion on a test, each item was reviewed and commented on by at least one other B&A employee. Defendant’s Exhibit N, App. T, contains charts linking each and every test item on all components of both the Fire Captain and Battalion Chief tests to both the major work behavior involved and the specific knowledge, skill or ability the item purports to test. The aforementioned “justifications”, as well as the charts showing linkages between job descriptions and KSAs, and between test items and major work behaviors and KSAs were completed before the tests were administered. These “justifications” were required to validate the tests.

Prior to the administration of the 2000 examinations, Barrett & Associates provided each candidate with a study guide manual for his or her use to prepare for the examination. The study guide manuals contained information about each exercise in the selection test battery (e.g. test component weights, reading lists, time limits, scoring and scoring dimensions), general test rules and procedures, general test taking strategies, sample test instructions, sample fire scene simulation exercises and test items, instructions for completing the computer-scored response sheets, and, for Battalion Chief candidates only, sample job knowledge test items.

Both examinations were administered in May, 2000. There was no cut-off, or passing, score set for any of the individual components on either examination, but each examination had an overall cutoff score of seventy (70) per cent. Barrett & Associates also instituted an “appeal” or “challenge” process whereby, candidates could contest specific test items and answers.

One hundred eighty-three (183) white candidates and one hundred twenty-four (124) black candidates completed all of the 2000 Fire Captain exam components. One hundred nineteen (119) of the white candidates and fifty- five (55) of the black candidates scored above the seventy per cent overall cutoff for the test and were listed in rank order by their scores on the Fire Captain's eligible list. The Fire Captain eligible list was in effect for approximately two years. During this time, eighteen (18) white candidates and four (4) black candidates were promoted to the position of Fire Captain.

Twenty-seven (27) white candidates and twenty- four (24) black candidates completed all four of the 2000 Battalion Chief examination components. Sixteen (16) white applicants and ten (10) black applicants scored above the seventy per cent overall cutoff for the test and were ranked according to their scores on the eligible list for the Battalion Chief position. The Battalion Chief eligible list was also in effect for approximately two years, during which time two (2) white candidates were promoted to the position of Battalion Chief. No black candidates were promoted.

B. 2004 Fire Captain and Battalion Chief Tests

In 2003, SHL USA ("SHL") was hired by the City to develop and validate promotional examinations for the positions of Fire Captain and Fire Battalion Chief. The work under the contract was performed by SHL's public safety division, formerly Landy, Jacobs and Associates. This division was headed by Dr. Rick Jacobs, who was formerly a principal in the firm of Landy, Jacobs and Associates. In January, 2004, E.B. Jacobs, a firm headed by Dr. Jacobs, purchased the public safety division back from SHL and completed the work yet to be performed under the contract with the City. Dr. Janet Echemendia, Ph.D., who has a Ph.D. in industrial organizational psychology, was the project director for the entire project.

The first stage of the project involved conducting job analyses for the two positions. Dr. Greg Loviscky, who has a Ph.D. in industrial organizational psychology and Jennifer LaBlanc, who has a bachelor's degree in industrial organizational psychology, were directly involved in conducting the job analyses. Preliminary task lists were prepared based on a library of tasks for the Fire Captain and Battalion Chief positions that the test consultants had developed over the past twenty years and revised after interviews and panel discussions with job incumbents for the respective positions. Task survey participants were later asked to describe how well these task lists described their jobs and 82.7% of the Fire Captains surveyed and 100% of the Battalion Chiefs surveyed responded that these task lists captured most or all of their jobs.

Detailed task surveys utilizing the task lists were completed by samples of Fire Captain incumbents and Battalion Chief incumbents. Each group of survey participants rated the importance of the tasks and duties on the lists that incumbents in their position would perform on "day one" of the job. These task and duty ratings were used to determine the context of the work simulation exercises (oral board exercises and the Battalion Chief written in-basket exercises). The information obtained during the initial interviews and observations were also used for this purpose.

Knowledge surveys and "Overall Duty/Source/Ability Surveys" were completed by a sample of Fire Captain incumbents. The Fire Captain incumbents who participated in the knowledge surveys rated the importance and the manner of use of 348 specific sources of a Fire Captain's knowledge. The knowledge sources were identified through the earlier interviews, observations and panel discussions as important knowledge sources that were used on the job. The knowledge source ratings were used to determine which specific knowledge sources would be retained on the reading list. The manner of use ratings were used to determine whether the

particular knowledge sources tested would be included on the closed-book test or the open-book test.

“Overall Duty/Ability Surveys” were completed by a sample of Battalion Chief incumbents. The Battalion Chief survey did not include questionnaires regarding knowledge sources because it was determined that a test directly assessing knowledge was not necessary for this position. The “Overall Duty/Source/Ability Survey” for the Fire Captains and the “Overall Duty/Ability Survey” for Battalion Chiefs required the survey participants to distribute 100 points across the duties set out in the job description for each position. There were twenty-nine duties for Fire Captains and seventeen duties for Battalion Chief. This distribution resulted in relative duty importance ratings.

The Fire Captain incumbents who completed the “Overall Duty/Source/Ability survey” linked fifteen general knowledge sources to the performance of the twenty-nine duties set out in the task list. These linkages, combined with the relative duty importance ratings allowed for the creation of a weighted mean for each overall knowledge source. This information was used to determine the percentage of items on each of the Fire Captain knowledge tests that would be devoted to each of the overall knowledge sources.

E.B. Jacobs relied on the importance ratings for the 348 specific knowledge sources to determine the number of items that would come from each specific section or chapter of the overall knowledge source. Fire Captains and Battalion Chiefs who completed the “Overall Duty/Source/Ability Survey” and the “Overall Duty/Ability Survey” were asked to determine which of twelve abilities traditionally utilized by public safety leaders were linked to the performance of each of the duties of each position. These linkages, combined with the relative

duty importance ratings allowed for the creation of a weighted mean for each ability. This information was used to identify the abilities to assess as part of the work simulation exercises.

The 2004 Fire Captain examination consisted of two components: 1) a multiple choice knowledge test, part of which was closed book and part of which was open book, and 2) two oral board exercises. The items on the knowledge test were written by independent subject matter experts who were directed to write a certain number of items from specific knowledge sources. These items were then reviewed by E.B. Jacobs staff and by a culture bias reviewer. The incident command oral board exercise for the Fire Captain exam was also created by independent subject matter experts and underwent the same review process as the items on the knowledge tests. The subordinate conference exercise was developed internally by E.B. Jacobs' public safety testing specialists and reviewed by the external subject matter experts.

The 2004 Battalion Chief examination also consisted of two components: 1) thirteen in-basket exercises requiring essay responses, and 2) three oral board exercises. The Battalion Chief in-basket test component was created by E.B. Jacobs' internal public safety staff and reviewed by external subject matter experts. The Battalion Chief incident command exercise was created in the same manner the incident command exercise for the Fire Captain examination was created. The subordinate conference and group conference exercises for the Battalion Chief examination were created in a similar fashion as the Fire Captain subordinate conference.

There was no cut-off score set for any of the individual components on either 2004 examination, but each examination had an overall cutoff score, or passing score, of seventy percent. One hundred and twenty-two (122) white applicants and seventy-five (75) black applicants completed the 2004 examination for Fire Captain. Fifty-two (52) white applicants and sixteen (16) black applicants passed and were placed on the eligible list.

Twenty-nine (29) white applicants and twenty-seven (27) black applicants took the 2004 examination for Battalion Chief. Twelve (12) white applicants and seven (7) black applicants passed the examination. Although there are several vacancies for each position, no promotions have been made from either list because of this litigation. Each list remains in effect.

C. Appointing Authority

In 2000, pursuant to Article XVIII of the Charter of the City of St. Louis, 1914, as amended, and Rule VII of the Rules of the Civil Service Commission of the City of St. Louis, the general rule was that, whenever a vacancy in a particular job classification occurred, the Director of Personnel was required to certify to the appointing authority the names of the top three individuals on the eligible list for that job classification. Thereafter, the appointing authority was required to interview the individuals certified and was permitted to select any one of them to fill the vacant position. A recent Charter amendment enacted since the 2004 Fire Captain and Battalion Chief lists became effective requires that the Director of Personnel certify the six highest-ranking individuals on the eligible list and permits the appointing authority to choose any one of those six certified individuals to fill the vacant position.

III. CONCLUSIONS OF LAW

Disparate impact claims under Title VII challenge “employment practices that are facially neutral in their treatment of different groups but that in fact fall more harshly on one group than another and cannot be justified by business necessity.” Mems v. City of St. Paul, 224 F.3d 735, 740 (8th Cir. 2000) (quoting Hazen Paper Co. v. Biggins, 507 U.S. 604, 609 (1993)). The current standard of proof in a discriminatory impact case was first enunciated in Griggs v. Duke Power Co., 401 U.S. 424, 432 (1971), then reinstated by the Civil Rights Act and codified in 42 U.S.C.

§2000e-2(k).⁴ Once a plaintiff has made a *prima facie* case of disparate impact from a selection procedure, the burden shifts to the defendant employer to show that the employment practice in question is "job-related for the position in question and consistent with business necessity . . ."

42 U.S.C. §2000e-2(k); Allen, et. al. v. Entergy Corp., et. al., 181 F.3d. 902 (8th Cir. 1999) (*citing* both Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975) and Griggs, supra.). Once the employer meets this burden, the plaintiff can still prevail if he can show that an alternative selection method exists that has a less disparate impact and would also serve the employer's legitimate business interest. Allen v. Entergy Corp., supra., citing Hawkins v. Anheuser-Busch, Inc., 697 F.2d. 810, 816 (8th Cir. 1983); *see also*, Albemarle Paper Co., 422 U.S. at 425. The FIRE plaintiffs stipulated that they were not pursuing this alternative selection method argument.

A. *Prima Facie* Case

In order to establish a *prima facie* case in a Title VII disparate impact case challenging the use of employment tests, a plaintiff must show that the tests in question select applicants for hire or promotion in a racial pattern significantly different from that of the pool of applicants.

Albemarle Paper Co. v. Moody, 422 U.S. at 425(*citing* McDonnell Douglas Corp. v. Green, 411

⁴The standard for the burden of proof in disparate impact cases has undergone significant revisions since it was first established in Griggs. Under Griggs, after the plaintiff establishes a *prima facie* case of disparate impact, the burden is on the employer to prove that the challenged procedure is related to safe and efficient job performance, and that an alternative less discriminatory practice or policy is not available. In 1989, the Supreme Court in Wards Cove Packing Co. v. Atonio, 490 U.S. 692 (1989), altered the burden of proof and held that after the plaintiff establishes a *prima facie* case of disparate impact, the defendant employer bore the burden of producing evidence of a legitimate business justification in defense of the challenged practice or policy. However, the burden of persuasion remained with the plaintiff. *See*, Bradley v. Pizzaco of Nebraska, 7 F.3d. 795, 797 (8th Cir. 1993)(pointing out the significant differences of Griggs and Wards Cove). In response to Wards Cove, Congress enacted the Civil Rights Act of 1991. One of its stated purposes was to codify the standard of proof enunciated in Griggs. The 1991 Act and its codification at 42 U.S. §2000e-2(k) make clear that the burden of production and the burden of persuasion regarding business justification rest with the defendant employer. Bradley, at 797; Allen v. Entergy Corp., 181 F.3d. 902 (8th Cir. 1999).

U.S. 792, 802 (1973)); *see also*, Ramirez, et. al. v. City of Omaha, et. al., 678 F.2d. 751, 753 (8th Cir. 1982)(affirming district court's application of Albemarle *prima facie* case standard to find that the plaintiffs failed to make a *prima facie* case of disparate impact in a Title VII challenge to employment tests).

In this case, the tests in question are the four promotional examinations challenges by Plaintiffs: the 2000 and 2004 Fire Captain tests and the 2000 and 2004 Battalion Chief tests.

It is not uncommon for plaintiffs to use statistical evidence to establish a *prima facie* case of disparate impact. *See*, Krauel v. Iowa Methodist Medical Center, 95 F.3d. 674, 681 (8th Cir. 1996)(*citing* Watson v. Ft. Worth Bank & Trust, 487 U.S. 977, 994 (1988) (the plaintiff must offer "statistical evidence of a kind and degree sufficient to show that the practice in question has caused the exclusion" of benefits because the plaintiff was within a protected class); Webb v. Derwinski, 868 F.Supp. 1184, 1190 (E.D.Mo. 1994) *aff'd* 68 F.2d. 479 (8th Cir. 1995)(unpublished decision); Leidig v. Honeywell, Inc., 850 F.Supp. 796, 801 (D.Minn. 1994). The EEOC has adopted what is known as the "Four-Fifths (4/5) Rule " for establishing a *prima facie* case of disparate impact. "A selection rate for any race, sex, on [sic] ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact." 29 C.F.R. §1607.4(D).

The EEOC has adopted the Uniform Guidelines to "assist employers . . . to comply with the requirements of Federal law prohibiting employment practices which discriminate on grounds of race . . ." Uniform Guidelines, Section 1A,B. In assessing adverse impact, the Uniform Guidelines reference *selection* rates, not *passing* rates. Uniform Guidelines, Section 4D. As a

result, the significant number is those actually selected, that is, actually *promoted*. See, e.g., Firefighters Institute for Racial Equality v. St. Louis, 616 F.2d 350, 356 (8th Cir. 1980) (where the Court applied the Four-Fifths Rule to promotions, rather than passing the exam); Alford v. City of Montgomery, 879 F.Supp. 1143, 1149 (M.D.Ala. 1995) (“Generally to succeed under [a disparate impact] theory, the plaintiff must establish a ‘meaningful statistical comparison’ between those employees eligible for promotion and those actually promoted” (emphasis added)); Nash v. Consolidated City of Jacksonville, Duval County, Florida, 895 F.Supp. 1536, 1542 (M.D.Fla. 1995). This is particularly true where, as here, the number of candidates who exceed the passing “score” greatly exceed the number of actual promotions. While the Uniform Guidelines reference the Four-Fifths Rule, the Uniform Guidelines at Question and Answer 20 establish that the Four-Fifths Rule should *not* be used when the number of selections is small:

. . . [A] difference of more than 20% in rates of selection may not provide a basis for finding adverse impact if the number of persons selected is very small. For example, if the employer selected three males and one female from an applicant pool of 20 males and 10 females, the 4/5ths rule would indicate adverse impact (selection rate for women is 10%; for men 15%; 10/15 or 66 2/3% is less than 80%), yet the number of selections is too small to warrant a determination of adverse impact. In these circumstances, the enforcement agency would not require validity evidence in the absence of additional information (such as selection rates for a longer period of time) indicating adverse impact.

44 Fed. Reg. 11, 996 (March 2, 1979) (emphasis added). The federal courts have also followed this rule. See, Mems v. City of St. Paul, 224 F.3d 735, 740 (8th Cir. 2000) (noting that, “However, even where the Four-Fifths Rule is violated, differences in the selection rates may not constitute an adverse impact where the sample size is too small to be statistically significant”); NAACP v. City of Mansfield, 866 F.2d 162, 168 (6th Cir. 1989) (“[T]his court has been cautious about giving too much weight to ‘four-fifths’ computations where the sample base is of less than significant proportions.”); Frazier v. Consolidated Rail Corp., 851 F.2d 1447, 1451 (D.C. Cir.

1988) (approving the district court's rejection of the 80% rule where the sample size was small); Fudge v. City of Providence Fire Dep't, 766 F.2d 650, 658-59 & n. 10 (1st Cir. 1985) (the four-fifths rule is not an accurate test of adverse impact when the sample size is small); Bridgeport Firebird Soc'y v. City of Bridgeport, 686 F.Supp. 53, 59 n. 6 (D. Conn. 1988) ("The 'four-fifths' rule may not be a valid indicator of adverse impact when applied to exceptionally small groups").

2000 Battalion Chief Exam

On the 2000 Battalion Chief examination, sixteen (16) white applicants and ten (10) black applicants scored above the seventy per cent overall cutoff for the test and were ranked according to their scores on the eligible list for the Battalion Chief position. Only two (2) candidates were promoted from the 2000 Battalion Chief exam. Plaintiffs' expert, Dr. Gary Renz, agreed with defendant's expert, Dr. Gerald Barrett, that the Uniform Guidelines caution against using the Four-Fifths Rule in dealing with numbers such as those promoted from the 2000 Battalion Chief exam. Uniform Guidelines at Question and Answer 21 is on point:

Generally, it is inappropriate to require validity evidence or to take enforcement action where the number of persons and the difference in selection rates are so small that the selection of one different person for one job would shift the result from adverse impact against one group to a situation in which that group has a higher selection rate than the other group.

While the two individuals who were promoted as a result of the 2000 Battalion Chief exam were both white, Dr. Renz admitted that if the race of one of these individuals was changed, then African Americans would have a higher selection rate than whites. As a result, pursuant to Question and Answer 21, "validity evidence" is not required, and "enforcement action" is not appropriate.

Dr. Renz also agreed that the selection rates for the 2000 Battalion Chief exam were not statistically significant in terms of racial impact. Indeed, Dr. Renz performed three statistical reviews — a Fischer Exact Test, a Chi-Square Value Test, and a Difference of Proportions Test — none of which showed any statistical significance in promotion rates.

As noted above, the critical inquiry is the existence or non-existence of a disparity in selection rates. Uniform Guidelines, Section 4D; Firefighters Institute for Racial Equality v. St. Louis, 616 F.2d at 356. However, it is noteworthy that the passing rate for the 2000 Battalion Chief exam also does not have a statistically significant adverse impact. Dr. Renz testified that, “The 2000 Battalion Chief examination did not have a statistically significant disparate impact against African American candidates in terms of the overall passing rates.” His conclusion that there was not a statistically significant difference in passing rates was confirmed by both a Chi-Square Value Test and a Difference of Proportions Test. Also, using a two-tailed test Dr. Renz concluded that there was no statistically significant difference in the rank order on the promotion eligibility list. Therefore, the FIRE plaintiffs failed to make a *prima facie* case of adverse impact in the 2000 Battalion Chief exam.

2000 Fire Captain Exam

One hundred eighty-three (183) white candidates and one hundred twenty-four (124) black candidates completed all of the 2000 Fire Captain exam components. One hundred nineteen (119) of the white candidates and fifty-five (55) of the black candidates scored above the seventy per cent overall cutoff for the test and were listed in rank order by their scores on the Fire Captain’s eligible list. The Fire Captain eligible list was in effect for approximately two years. During this time, eighteen (18) white candidates and four (4) black candidates were promoted to the position of Fire Captain.

The City's expert, Dr. Barrett, opined that there was no adverse impact when the number of promotions for each race were compared with the numbers from each race who passed. He testified that using the number of passing applicants as the benchmark, rather than using the number of total candidates who took the test, was justified because the comparison must involve those candidates who were minimally qualified. Those who did not pass lacked this minimum qualification because they could not be on the eligibility list and could not be promoted without having a passing score. In his expert report, Dr. Barrett noted that if the number of blacks who were promoted was increased slightly, the black promotion rate was slightly more than the white promotion rate. Therefore, Dr. Barrett testified, the Four-Fifths Rule was inapplicable. Dr. Barrett also performed a Fischer's Exact Test to determine whether there was a statistically significant impact based on race. He concluded there was not.

The FIRE plaintiffs' expert, Dr. Renz, testified that there was a violation of the Four-Fifths Rule with respect to the 2000 Fire Captain's examination when looked at from the point of view of passing rates. However, as discussed above, the Uniform Guidelines reference *selection* rates, not *passing* rates. Uniform Guidelines, Section 4D; Firefighters Institute for Racial Equality v. St. Louis, 616 F.2d at 356. With regard to promotion rates, Dr. Renz also testified that the 2000 Fire Captain's examination violated the Four-Fifths Rule. In performing this calculation, Dr. Renz used the total number of exam candidates of each race as his benchmark, rather than the total number of candidates from each race who scored above the seventy percent cutoff. The Eighth Circuit has previously used the total applicant pool, rather than the number of candidates who passed the minimum threshold, to calculate statistical significance in the determination of adverse impact. Id. at 356, n. 6. Therefore, I find that Dr. Renz has utilized an accurate benchmark and therefore the Four-Fifths Rule is applicable to the

2000 Fire Captain exam. Accordingly, the FIRE plaintiffs have made a *prima facie* case of adverse impact in the 2000 Fire Captain exam.

2004 Battalion Chief and Fire Captain Exams

With respect to the 2004 tests, it is not possible based on a traditional analysis to determine whether there has been an adverse impact under the Four-Fifths Rule because by definition, adverse impact under the Four-Fifths Rule requires the *selection rate* of a protected class be less than four-fifths (4/5) (or eighty percent) of the *selection rate* of the most selected group. 29 C.F.R. §1607.4(D). Because no one has yet been selected from the list generated by the 2004 Battalion Chief and Fire Captain tests, there cannot yet be a determination of adverse impact. However, it appears that the City has purposefully waited to make any promotions off of the 2004 lists pending the outcome of this litigation. I will therefore continue with the analysis as though an adverse impact had been found in the 2004 examinations.

Conclusion

After reviewing the evidence submitted, I conclude that the statistics presented here do not establish a *prima facie* case of disparate impact discrimination on the 2000 Battalion Chief examination. The statistics do establish a *prima facie* case of disparate impact discrimination on the 2000 Fire Captain examination. With respect to the 2004 Fire Captain and Battalion Chief examinations, because no one has yet been selected from the list generated by the 2004 Battalion Chief and Fire Captain tests, there cannot yet be a determination of adverse impact. However, because it appears that the City has purposefully waited to make any promotions off of the 2004 lists pending the outcome of this litigation, I will continue with my analysis as though an adverse impact had been found in the 2004 examinations.

B. Validity

Because the FIRE plaintiffs have met their burden on establishing a *prima facie* case of disparate impact on the 2000 Fire Captain examination, the burden shifts to the City to prove that its test was “job-related.” At trial, the City produced overwhelming evidence tending to prove that the 2000 Fire Captain examination was “job-related for the position in question and consistent with business necessity” as required by 42 U.S.C. § 2000e(k). Additionally, even had the FIRE plaintiffs met their burden of establishing a *prima facie* case of disparate impact on the 2000 and 2004 Battalion Chief examinations and the 2004 Fire Captain examination, the City also produced overwhelming evidence at trial tending to prove that each of those three examinations were “job-related for the position in question and consistent with business necessity” as required by 42 U.S.C. § 2000e(k).

“Business necessity” is defined in Section 105(b) of the 1991 Civil Rights Act, 101 P.L. 166, exclusively by reference to an interpretive memorandum set out in Vol. 137 Congressional Record S 15276, daily ed. Oct. 25, 1991, which states that the terms “business necessity” and “job related” are intended to reflect the concepts enunciated by the Supreme Court in Griggs v. Duke Power Co., 401 U.S. 424 (1971) and the other Supreme Court decisions prior to Wards Cove Packing v. Atonio, 490 U.S. 642 (1989). In order to prove business necessity, an employer must show that its selection criteria bear “a manifest relationship to the employment in question.” Griggs v. Duke Power Co., 401 U.S. at 432. The employer must also demonstrate that the employment practice significantly serves legitimate employment goals. See New York City Transit Authority v. Beazer, 440 U.S. 568, 587 n.31 (1979). However, the employer is not required to show that those employment goals “require” the employment practice. See Id. Similar definitions have been adopted by the Eighth Circuit in Firefighters Institute for Racial

Equality v. City of St. Louis, 220 F.3d 898, 903 (8th Cir. 2000) (“manifest relationship to the employment in question,” “related to safe and efficient job performance”) and Allen v. Entergy, 181 F.3d 902, 904 (8th Cir. 1999) (challenged practice must be “related to safe and efficient job performance”).

2000 Fire Captain and Battalion Chief Exams

The testimony at trial described Barrett & Associates’ methods used in developing, administering and scoring the promotional examination. The testimony also explained how the questions for the written test, and other examination components, were developed so that the content of the questions would be tied to critical and frequent work tasks of a Fire Captain and Battalion Chief.

The testimony at trial showed that the test complies with the EEOC’s Uniform Guidelines on Employment Selection Procedures, 29 C.F.R. § 1607.5. "The EEOC has issued 'Guidelines' for employers seeking to determine, through professional validation studies, whether their employment tests are job related." Albemarle Paper Co., 422 U.S. at 430-31. Although these guidelines have not been promulgated pursuant to formal congressional procedures, *see* Albemarle Paper Co., 422 U.S. at 431, they nevertheless are "entitled to great deference" as "[t]he administrative interpretation of the Act by the enforcing agency, " Griggs v. Duke Power Co., 401 U.S. at 433-34; *see also* Allen v. Entergy Corp., *supra*.(quoting both Albemarle and Griggs); Firefighters Institute for Racial Equality, et. al. v. City of St. Louis, 549 F.2d. 506, 510 (8th Cir. 1977)(application of the EEOC’s Guidelines in resolving issue of job-relatedness). The record contains charts linking each and every test item on all components of both the Fire Captain and Battalion Chief tests to both the major work behavior involved and the specific knowledge, skill or ability the item purports to test.

The only evidence offered by the FIRE plaintiffs to rebut the City's evidence that the tests were valid and job-related was the expert testimony of Dr. Renz. Dr. Renz admitted on cross-examination that he did not detect anything that was not job-related in the component parts of the 2000 examinations.

Dr. Renz's major criticism of the 2000 Fire Captain and Battalion Chief tests was that evidence of the job analysis was not adequate. Dr. Arthur and Dr. Barrett both testified that Barrett & Associates did not use a "questionnaire" approach. Rather, Dr. Arthur testified that Dr. Barrett, and Dr. Michael Polomsky and he used a "semi- structured interview approach" in which they gathered information from job incumbents by inquiring about such things as what tasks they performed in a typical day, followed by questions eliciting information as to how much supervision or discretion was involved in each task, how frequently each task was performed and how critical the task was to the job overall. Not every interviewee was asked the same questions; rather, the inquiries would follow up on the responses received from each interviewee. The interviewers also reviewed "archival material," including previous job descriptions, job announcements and specifications, in performing the job analyses.

Because no standardized questionnaire was used, there was no "paper trail" of the type Dr. Renz testified he believed was necessary. Instead, at the end of each day of interviews, the interviewers reported their findings, which were based on notes taken at the interviews, to Dr. Arthur who consolidated the information into what eventually emerged as comprehensive and detailed job descriptions for both the Fire Captain and Battalion Chief positions. In addition to listing the major work behaviors of each position, the job descriptions also listed the behaviors in order of importance. These rankings were based on the responses to the semi-structured interview questions. Moreover, the job descriptions also listed the knowledge, skills and abilities

(“KSAs”) associated with the performance of each major work behavior, based on both interview responses and the professional judgment of the interviewers.

Finally, when the interviews were completed and draft job descriptions produced, the same incumbents who had been interviewed were asked to review the draft job description and return it with their comments and proposed changes. Dr. Arthur testified that he received comments on the draft job descriptions, which were then incorporated into the final job descriptions. The result of the “consensus” process was that the final job descriptions fairly represented the jobs in question.

Dr. Renz’s assertion that he had no way of knowing whether the final job description actually reflected the job it purported to define is not supported by the evidence of how the job descriptions were actually drafted. The job incumbents, themselves, had already made it clear that the job descriptions accurately reflected their jobs. Dr. Barrett and Dr. Arthur both testified that there are a number of methods of conducting a job analysis. Dr. Arthur noted in his testimony that the job analysis was a *process* which results in the job description. The fact that the semi-structured interview method utilized in developing the job descriptions for the 2000 examinations did not produce the precise type of documentary linkage Dr. Renz testified he prefers did not make the resulting job descriptions any less valid.

Dr. Renz also raised the issue that he felt there was an uneven coverage of major work behaviors in the 2000 examinations. Dr. Renz noted that the vast majority of the questions in the job knowledge test referred to Major Work Behaviors I and II (suppressing fires and responding to emergency calls, respectively). Dr. Barrett responded to this criticism by noting simply that those two major work behaviors were the “main responsibilities” for both the Fire Captain and Battalion Chief positions and that the tests, therefore, focused on them because they were the

most important work behaviors. Dr. Jacobs, responding to a similar criticism of the 2004 examinations by Dr. Renz, discussed the issue in terms of errors of commission and omission and essentially suggested that it is a greater error to include something in a test that has little to do with the job, than it is to omit questions on matters that are related to the job. In other words, Dr. Barrett and Dr. Jacobs testified that focusing test questions on the major work behaviors of fire suppression and emergency response is far more likely to predict success on the job as a Fire Captain or Battalion Chief than focusing on a lower-ranked behavior such as the upkeep of buildings and grounds (Fire Captain's Major Work Behavior VII).

The balance of Dr. Renz's criticisms of the 2000 examinations stemmed primarily from his claims that not enough documentation was provided to him to make certain determinations and that, therefore, the tests were invalid. However, nothing in the Guidelines says that everything a reviewer needs to determine validity has to be included in a validity study, or report. The City argues that the validity studies submitted by Dr. Barrett provided more than enough information and documentation for Dr. Renz to make substantive determinations with respect to the validity of the 2000 examinations. It is noteworthy that, although Dr. Renz complains repeatedly in his reports that there is not "enough" information, he never indicates how much, or what, "enough" would be. More importantly, he entirely fails to cite to any standard which defines what documentation is "enough" to determine validity. As a result, Dr. Renz has made his determinations of invalidity on a purely subjective basis, because what is "not enough" information for him might very well be "more than enough" for a different reviewer. In other words, while he challenges the methodology of the test drafters, he wholly fails to set forth his own methodology or standards for rendering his opinions. Dr. Renz's "say so" without a more rigorous standard or methodology is not persuasive.

Far from requiring that all pertinent validity documentation be included in a report, the Standards actually recognize that it may not be practical to do so. In fact, placed in context, page 67 of the Standard relied upon by Dr. Renz states:

Test documents need to include enough information to allow test users and reviewers to determine the appropriateness of the test for its intended purpose. References to other materials that provide more details about research by the publisher or independent investigators should be cited and should be readily obtainable by the test user or reviewer. This supplemental material can be provided in any of a variety of published or unpublished forms; when demand is likely to be low, it may be maintained in archival form, including electronic storage.

Therefore, Dr. Barrett's validity studies, taken together with his testimony and that of his colleagues, satisfies the Standards.

Dr. Renz suggested that Barrett & Associates should have utilized concurrent criterion-related validity evidence in its efforts to show that the examinations were job-related. Content validity evidence and criterion-related validity evidence are two different types of validity evidence that can be utilized to show "job relatedness." Content validity evidence is evidence that links the content of the examination to the KSAs construct that underlie the performance of tasks that are a part of the job at issue. This is the type of validity evidence relied upon in validating the 2000 examinations. A job analysis was conducted to identify the important tasks and duties of the job and then the KSAs underlying these tasks and duties were identified. Lastly, tests were constructed to measure these KSAs. Content validity evidence is a recognized source of validity evidence under the Guidelines. Dr. Barrett testified that most employment tests in both the public and private sectors are content validated.

Criterion-related validity evidence is empirical data linking test scores (predictor information) to job performance (criterion information). Dr. Jacobs testified that a concurrent criterion-related validity study is where the examination is administered to a group of job

incumbents prior to administering it to job candidates or administering it at the same time and then measuring the job performance of the incumbents to determine whether their test scores are correlated to their job performance. Dr. Renz testified that concurrent criterion-related validity evidence would have been a more adequate indicator of validity, even though he also testified that he knew of *no instances* where criterion-related validity evidence has been utilized to establish the job-relatedness of promotional testing. Dr. Barrett testified that he had never seen such evidence obtained in promotional testing. Dr. Barrett testified that there are serious security concerns with administering the tests to job incumbents before they were administered to the candidates. The security concerns related to the Battalion Chief examination would be compounded because the incumbents who would be tested for the Fire Captain's concurrent validity study would also be candidates for the Battalion Chief examination, which was a similar examination. Furthermore, in regard to any criterion-related validity study of the Battalion Chief examination, the number of incumbents and candidates would likely be too small for the study to have any statistical significance. Dr. Barrett and Dr. Jacobs testified that such a validation study for either the Fire Captain examination or the Battalion Chief examination would present the problem of finding incumbents willing to take the examinations, which spanned multiple days, and incumbents motivated to prepare and study for the examination. Furthermore, such a study faces the obstacle of finding superiors willing to rate and rank incumbents in their job performance and stand by those ratings and rankings. Dr. Barrett testified that this is a highly unlikely proposition when the likelihood of ensuing litigation is high.

Dr. Renz's final criticism of the 2000 examinations was that they should have been reviewed by personnel in the St. Louis Fire Department prior to administration. Dr. Renz testified that he found it "hard to believe" that "nobody in the fire department could be trusted

enough to review the tests.” Despite Dr. Renz’s view, Fire Chief Sherman George testified that he did not want to participate in such a review. Dr. Jacobs testified that although the infeasibility of a preview of the tests by the St. Louis Fire Department is unfortunate, the lack of such a review does not render the tests invalid.

2004 Fire Captain and Battalion Chief Examinations

The testimony at trial described E.B. Jacob’s methods used in developing, administering and scoring the promotional examinations. The testimony also explained how the questions for the written test and other examination components were developed to ensure that the content of the questions would be tied to the critical and frequent work tasks of Fire Captains and Battalion Chiefs.

I find that the testimony at trial showed that the tests comply with the EEOC’s Uniform Guidelines on Employment Selection Procedures, 29 C.F.R. § 1607.5. The record contains charts linking each and every test item on all components of both the Fire Captain and Battalion Chief tests to both the major work behavior involved and the specific knowledge, skill or ability the item is designed to test.

Again, Plaintiffs’ only evidence offered to rebut the City’s evidence that the tests were valid and job-related was the expert testimony of Dr. Renz. Dr. Renz admitted on cross-examination that he did not detect anything that was not job-related in the component parts of the 2004 examinations.

E.B. Jacobs relied upon content validity evidence to support the validity and job-relatedness of the tests administered in 2004. Content validation relies upon “data showing that the content of the selection procedure is representative of important aspects of performance on the job”. 29 C.F.R. §1607.5 (B). In her testimony, Dr. Echemendia described in detail the steps

taken to create an accurate job description for each position at issue. She described the steps taken to discern the abilities, and in the case of the Fire Captain position, the knowledge sources, important to job performance and the steps taken to construct test components that measure these constructs. The data supports these linkages.

The data produced at trial shows that the task lists accurately describe the Fire Captain and the Battalion Chief jobs. The data establishes the most important duties. The data identifies the most important specific knowledge sources for the Fire Captain position and the manner in which these knowledge sources are used. The data also identifies the most important overall knowledge sources for the Fire Captain position and the most important abilities for both the Fire Captain position and the Battalion Chief position. This data was then used to construct the examination components.

Dr. Jabobs testified that the data shows that the Fire Captain knowledge tests were a reliable measure of the knowledge identified as important to the jobs at issue. With respect to the work simulation exercises, E.B. Jacobs utilized experienced fire fighting personnel from fire departments outside of the State of Missouri to assess candidate performance. The assessors were a racially diverse group of command- level fire fighters from both medium- size and large fire departments across the country. Dr. Echemendia testified that the assessors underwent extensive training in regard to understanding and recognizing the abilities being tested and in regard to scoring those abilities. Twelve panels of three assessors each rated the oral board exercises for the Fire Captain examination. Nine of these same assessors, who held the rank of Battalion Chief or higher, rated the Battalion Chief oral board exercises. Four panels of two assessors each rated the Battalion Chief in-basket exercises.

Dr. Jacobs testified that the internal consistency, or inter-rater reliability, of both the Fire Captain and Battalion Chief oral board exercises and the Battalion Chief in-basket exercises was very high. This means that the scores of the raters on the same panels were highly correlated – the assessors agreed about what the candidate was doing and the score he or she should receive.

Dr. Jacobs testified that while the data did show a statistically significant difference in candidate scores on the oral board exercises for both positions as a function of panel assignment, this not an unusual occurrence when utilizing assessor panels and can be addressed by standardizing the scores by rescaling the mean dimension ratings within each panel. Monitoring the assessors during the scoring process, as suggested by Dr. Renz, would have been a “bad idea” because it would have disrupted the scoring process and would have given candidates reason to question the scoring process.

At trial, Dr. Renz testified that E.B. Jacobs used a sample, rather than a census, of job incumbents to complete the survey questionnaires. It was not clear whether this was a point of criticism or simply a comment on the process. However, in his report, Dr. Renz was clearly critical of E.B. Jacobs’ use of a sample and suggested that a census should have been used. This criticism was nothing more than nitpicking because this opinion does not go to the validity of the tests. Furthermore, Dr. Renz admits that none of the authorities that he relies on require that a job analysis be based upon a census and admits that sampling is acceptable. The Principles note only that “studying multiple incumbents may be necessary to understand differences in work complexity, work content, work environment, job behaviors, or worker KSAOs as a function of shift, location, variations in how work is performed and other factors....”

Thirteen out of eighteen incumbent Battalion Chiefs participated in the job survey. Dr. Jacobs testified that this was a representative cross-section and the additional five Battalion Chiefs would not have changed the results. Twenty-nine incumbent Fire Captains participated in the first part of the job survey (task survey) for Fire Captains and another thirty participated in all or part of the second part of the job survey. Dr. Jacobs testified that this was a stable number given that E.B. Jacobs obtained a sample that was representative of the Fire Captains in the department with respect to shift and assignment, as well as race and gender.

Dr. Renz was also critical of E.B. Jacobs' use of relative duty importance ratings to create the weighted knowledge source ratings and the weighted ability ratings. He testified that assigning a percentage to each of the duties (29 for Fire Captains and 17 for Battalion Chiefs) was too cognitively difficult a task for the survey participants and he therefore assumed that the results were not reliable. Further, in his report, Dr. Renz asserted that the survey participants did not have the job descriptions which listed the tasks associated with each duty when they completed these ratings.

Dr. Jacobs testified that he and another industrial organizational psychologist started using relative importance of duties ratings in 1985 and that the incumbents who completed the surveys would indeed have been capable of understanding the concept of percentages and adding up 100 points. Further, the survey incumbents did, in fact, have the job descriptions when they completed this part of the survey.

In his testimony, Dr. Renz also noted that E.B. Jacobs "brought these twelve different abilities to the raters in the job analysis phase . . ." In his report, he criticized E.B. Jacobs for not having the survey participants identify abilities or requesting the addition of any duties not already identified. Dr. Jacobs testified that the twelve abilities are a taxonomy, or scheme of

classification, of leadership abilities that E.B. Jacobs has developed over the years from studying the jobs of public safety leaders. Dr. Jacobs testified that there could be no systematic collection of data if survey participants were asked to create the list of abilities. The participants would likely label abilities differently, although meaning the same thing. Furthermore, there was no need to ask participants to identify additional abilities. Dr Jacobs testified that over the years, such a question has elicited very little, or no, response.

Dr. Renz criticized E.B. Jacobs for allegedly failing to sufficiently document the reasons for testing the abilities tested. In E.B. Jacobs' final report, it was noted that E.B. Jacobs sought to test "the abilities found to be important and most amenable to testing". The Fire Captain oral board exercises tested six of the nine highest rated abilities. The Battalion Chief in-basket and oral board exercises, together, tested seven of the ten highest rated abilities. Dr. Jacobs explained why some of the abilities, although highly ranked, were not tested. He testified that behavioral flexibility, and creativity and originality are very difficult abilities to measure and that his firm had attempted to measure behavioral flexibility in the past and had found it very difficult to do so and obtain reliable judgments. He further explained that a test measuring memory would be in a multiple choice format that would not be appropriate in the work simulation setting. Dr. Jacobs also testified that it is not necessary to test all of the abilities that are important to the job and that the abilities are all significantly correlated with one another so that by tapping a subset, a good number are covered. Significantly, although critical of E.B. Jacobs for not testing all the highest rated abilities, Dr. Renz offered no opinion as to how a test consultant would test those abilities not tested by E.B. Jacobs.

Dr. Renz also was critical of the fact that scores for the same abilities across different exercises were not highly correlated with one another, whereas the scores for different abilities

tested in connection with each exercise were highly correlated with one another. Based on these correlations as well as on a factor analysis he conducted that purportedly shows that the only factor being measured was performance on the exercise, Dr. Renz concluded that the assessors rated the candidates' performance on each exercise and not the underlying abilities. Dr. Jacobs testified that the fallacy of Dr. Renz's opinion is that it is based on the assumptions that the abilities are unrelated to one another and that a candidate who does well in one work situation will do just as well in another work situation. As Dr. Jacobs explained, these are not valid assumptions. This is why more than one exercise measuring the same abilities is administered. Furthermore, Dr. Jacobs testified that Dr. Renz's factor analysis is fundamentally flawed in that it assumes that the underlying abilities are not correlated, when indeed they are, and that he had an insufficient number of subjects for each variable to conduct such a statistical analysis.

One of Dr. Renz's biggest complaints about the 2004 examinations is that he was unable to determine from the final report how E.B. Jacobs moved from the job analyses to the tests themselves. Both Dr. Echemendia and Dr. Jacobs testified at length about the development of the selection processes and, specifically, how the data obtained from the job analyses were utilized to construct the test components. The final report did provide an overview of the development of both selection procedures which, supplemented with the testimony of both Dr. Echemenda and Dr. Jacobs, leads me to conclude that the content of the tests is linked to the important duties and tasks of the jobs of Fire Captain and Battalion Chief.

As with his criticisms of the 2000 examinations, the remainder of Dr. Renz's criticisms of the 2004 examinations stemmed primarily from his claims that not enough documentation was provided for him to make certain determinations and that, therefore, the tests were invalid. As

discussed above, E.B. Jacob's expert report, taken together with Dr. Jacobs' testimony and that of his colleagues, satisfies the Standards.

Dr. Renz also suggested that E.B. Jacobs should have utilized concurrent criterion-related validity evidence, in its efforts to show that the examinations were job related. This reasoning fails for the reasons I discussed above with the 2000 examinations.

Dr. Renz's final criticism of the 2004 examinations was that they should have been reviewed by personnel in the St. Louis Fire Department prior to administration. As I discussed earlier, Dr. Renz testified that he found it "hard to believe" that "nobody in the fire department could be trusted enough to review the tests." But again I note that even Fire Chief Sherman George testified that he does to want to participate in such a review. Dr. Jacobs testified that although the infeasibility of a preview of the tests by the St. Louis Fire Department is unfortunate, the lack of such a review does not render the tests invalid.

Conclusion

I find that the City of St. Louis has presented sufficient evidence to demonstrate that the 2000 and 2004 Fire Captain and Battalion Chief examinations were job related for the positions in question and consistent with business necessity. I therefore find that the Fire Captain and Battalion Chief examinations administered in 2000 and 2004 were valid examinations.


IV. CONCLUSION

The 2000 examination for the position of Battalion Chief did not result in an adverse impact on African-Americans. The 2000 examination for the position of Fire Captain did result in an adverse impact on African-Americans. Because no one has yet been selected from the list generated by the 2004 Battalion Chief and Fire Captain tests, there cannot yet be a determination of adverse impact as to these two examinations.

Defendant City of St. Louis has presented sufficient evidence to demonstrate that the 2000 and 2004 Fire Captain and Battalion Chief examinations were job related for the positions in question and consistent with business necessity. The City of St. Louis has therefore met its burden of production required to demonstrate that the Fire Captain and Battalion Chief examinations administered in 2000 and 2004 were valid examinations.

The FIRE plaintiffs did not produce any evidence of alternative selection procedures of which the City of St. Louis was apprised which would have been as substantially valid as the 2000 and 2004 Fire Captain and Battalion Chief examinations, but with lesser adverse impact. The FIRE plaintiffs have therefore failed to meet the burden of proof required for a finding that the 2000 and 2004 Fire Captain and Battalion Chief examinations were discriminatory by reason of adverse impact on African-Americans.

As a result, I will enter judgment in favor of Defendant, the City of St. Louis.



RODNEY W. SIPPEL
UNITED STATES DISTRICT JUDGE

Dated this 25th day of May, 2007.